

***IN THE UNITED STATES PATENT AND TRADEMARK OFFICE***

Applicant: Kiyoshi SHINOMIYA  
Title: INTER-COMPUTER DATA TRANSFER METHOD AND INTER-COMPUTER  
NETWORK SYSTEM  
Appl. No.: Unassigned  
Filing Date: 02/20/2004  
Examiner: Unassigned  
Art Unit: Unassigned

**CLAIM FOR CONVENTION PRIORITY**

Commissioner for Patents  
PO Box 1450  
Alexandria, Virginia 22313-1450

Sir:

The benefit of the filing date of the following prior foreign application filed in the following foreign country is hereby requested, and the right of priority provided in 35 U.S.C. § 119 is hereby claimed.

In support of this claim, filed herewith is a certified copy of said original foreign application:

- Japanese Patent Application No. 2003-044198 filed 02/21/2003.

Respectfully submitted,

*Michael M. A. Blumenthal*

*Reg No. 34,717*

By \_\_\_\_\_

Date: February 20, 2004

FOLEY & LARDNER  
Customer Number: 22428  
Telephone: (202) 672-5407  
Facsimile: (202) 672-5399

*for*

David A. Blumenthal  
Attorney for Applicant  
Registration No. 26,257

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日                      2 0 0 3 年    2 月 2 1 日  
Date of Application:

出 願 番 号                      特 願 2 0 0 3 - 0 4 4 1 9 8  
Application Number:

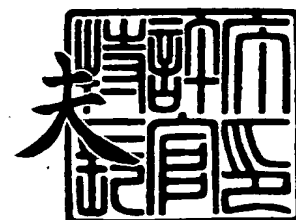
[ST. 10/C] :                      [ J P 2 0 0 3 - 0 4 4 1 9 8 ]

出      願      人                      日 本 電 気 株 式 有 限 公 司  
Applicant(s):

2 0 0 4 年    1 月 1 3 日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

今 井 康 夫



【書類名】 特許願

【整理番号】 66206635

【提出日】 平成15年 2月21日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 15/16

【発明者】

【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

【氏名】 四宮 潔

【特許出願人】

【識別番号】 000004237

【氏名又は名称】 日本電気株式会社

【代理人】

【識別番号】 100088328

【弁理士】

【氏名又は名称】 金田 暢之

【電話番号】 03-3585-1882

【選任した代理人】

【識別番号】 100106297

【弁理士】

【氏名又は名称】 伊藤 克博

【選任した代理人】

【識別番号】 100106138

【弁理士】

【氏名又は名称】 石橋 政幸

【手数料の表示】

【予納台帳番号】 089681

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9710078

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 データ転送方法およびこれによるネットワークシステム

【特許請求の範囲】

【請求項 1】 複数の計算機がパケット交換網を介して接続されるネットワークシステムで行なわれるデータ転送方法において、

前記複数の計算機のそれぞれに、

主記憶装置と、

通信処理命令を発行するプロセッサと、

前記プロセッサからの通信命令を処理し、前記パケット交換網を介して他の計算機と通信を行う通信装置とを設け、

前記通信装置には、

交換網に通信を送信する送信部と交換網から通信を受信する受信部それぞれが T L B エントリを複数保持する T L B を設け、

前記プロセッサは送り元の計算機の情報を含む前記通信命令を発行することとし、前記受信部は、使用する T L B エントリを送り元の計算機に応じて決定させることを特徴とするデータ転送方法。

【請求項 2】 請求項 1 記載のデータ転送方法において、

送信部は、送り元の計算機の情報を含む通信パケットを生成して他の計算機へ送信することを特徴とするデータ転送方法。

【請求項 3】 請求項 1 または請求項 2 に記載のデータ転送方法において、

通信装置は、送り元の計算機から同一宛先計算機に対する同時に処理する通信命令の数を制限して処理を行なうことを特徴とするデータ転送方法。

【請求項 4】 請求項 2 または請求項 3 に記載のデータ転送方法において、

通信装置は、送り元の計算機における処理命令の識別番号を通信パケットに付与することとし、受信部は、前記処理命令の識別番号が同じである通信パケットについては使用する T L B エントリを同じとすることを特徴とするデータ転送方法。

【請求項 5】 複数の計算機がパケット交換網を介して接続されるネットワークシステムにおいて、

前記複数の計算機のそれぞれは、  
主記憶装置と、  
通信処理命令を発行するプロセッサと、  
前記プロセッサからの通信命令を処理し、前記パケット交換網を介して他の計算機と通信を行う通信装置とを具備し、  
前記通信装置は、  
交換網に通信を送信する送信部と交換網から通信を受信する受信部それぞれが T L B エントリを複数保持する T L B を具備し、  
前記プロセッサは送り元の計算機の情報を含む前記通信命令を発行し、前記受信部は、使用する T L B エントリを送り元の計算機に応じて決定することを特徴とするネットワークシステム。

【請求項 6】 請求項 5 記載のネットワークシステムにおいて、  
送信部は、送り元の計算機の情報を含む通信パケットを生成して他の計算機へ送信することを特徴とするネットワークシステム。

【請求項 7】 請求項 5 または請求項 6 に記載のネットワークシステムにおいて、

通信装置は、送り元の計算機から同一宛先計算機に対する同時に処理する通信命令の数を制限して処理を行なうことを特徴とするネットワークシステム。

【請求項 8】 請求項 6 または請求項 7 に記載のネットワークシステムにおいて、

通信装置は、送り元の計算機における処理命令の識別番号を通信パケットに付与し、受信部は、前記処理命令の識別番号が同じである通信パケットについては使用する T L B エントリを同じとすることを特徴とするネットワークシステム。

#### 【発明の詳細な説明】

#### 【0001】

#### 【発明の属する技術分野】

本発明は、複数の計算機が互に通信しながら各計算機上のジョブの処理を並列に実行するネットワークシステムと、前記計算機システムにおけるデータ転送方法に関する。

**【0002】****【従来の技術】**

システム・エリア・ネットワークの普及により、複数の計算機を高速ネットワークにより接続し、各計算機が互いに通信しながら各計算機上のジョブの処理を並列に実行する分散コンピューティング方式が増加してきている。

**【0003】**

複数の計算機による並列処理により、一つの計算機では実現不可能な大規模演算処理を実行することが可能となる。また、複数の計算機で通信を行いつつ演算を行うプログラム作成用の標準的なプログラム言語であるHPF(High Performance Fortran)あるいはライブラリであるMPI(Message Passing Interface)も開発され、並列処理は益々広範な領域で使用されるものとなっており、前記の機能を有する計算機が増加してきている。

**【0004】**

複数の計算機をネットワークで接続して相互通信による並列処理を実行すると、実際の演算を行う演算処理と、他の計算機との通信を行う通信処理とが最低限必要となる。

**【0005】**

通信処理を行うにあたり、プロセッサが通信処理を行うのでは、通信処理のために演算能力が低下する。プロセッサの演算能力を可能な限り演算処理に使用するため、通信処理をプロセッサから取り除き、通信装置を実装することがある。

**【0006】**

通信装置は、プロセッサからの通信命令を処理し、通信命令により指定された送り元のデータを指定された宛先に転送する。通信命令においては、送り元データの主記憶上のアドレスである送り元アドレス、転送するデータの長さである転送長、送り先の計算機を指定することにより実施されることが一般的である。

**【0007】**

通信命令において、上記の送り元アドレス、転送長、送り先計算機を指定するのはSend命令と呼ばれ、上記に加えて送り先の主記憶アドレスである送り先ア

ドレスを指定するものはRemote DMA Write命令と呼ばれる。

#### 【0008】

通信命令にて指定される主記憶上のアドレスを物理アドレスとすると、通信処理装置に通信命令を送付するプロセスは物理アドレスを知ることができる特権プロセスに限られる。非特権プロセスであるユーザ・レベル・プロセスが通信処理を行う場合には、特権プロセスに通信を要求し、プロセス・スイッチにより特権プロセスが論理アドレスを物理アドレスに変換する処理を代理で実行する。このようなプロセス・スイッチ処理はユーザプロセスから見ると非常に長い時間がかかるため、プロセス・スイッチ処理が性能低下要因の一つとなっている。

#### 【0009】

上記のプロセス・スイッチ処理による性能低下を回避するために、通信装置が論理アドレスを物理アドレスに変換するアドレス変換機構を実装し、ユーザプロセスが特権プロセスを介することなく通信装置への通信命令発行を可能とすることをユーザ・レベル通信と呼び、低いレイテンシでの高性能の通信が要求されるネットワーク装置において実現されている。

#### 【0010】

アドレス変換機構を実装するにあたり、アドレス変換表を通信装置内に全て実装する方式、あるいはアドレス変換表を通信装置内に実装できない場合にはアドレス変換表を通信装置外部のメモリに実装し、通信装置にはその一部をTLB (Translation Look aside Buffer)として実装する方式が採用されている。

#### 【0011】

一般にTLBを実装する方式としては、以下の3方式が挙げられる。

#### 【0012】

(1) 変換対象論理アドレスを用いて、全TLBエントリの論理アドレスの一部を比較し、一致したTLBエントリをアドレス変換に用いるフル・アソシアティブ方式

(2) 変換対象論理アドレスの一部のビットを変換に使用するTLBエントリのアドレスとするダイレクト・マップ方式

(3) 変換対象論理アドレスの一部のビットにより全TLBから一部のTLB



エントリのセットを選択し、上記選択されたセット内の全TLBエントリ内の論理アドレスと変換対象論理アドレスを比較するセット・アソシアティブ方式

上記のいずれの方式においても、TLBエントリの決定には変換対象論理アドレスを用いている。

#### 【0013】

TLBを実装する方式を採用した場合には、TLBのヒット率が転送性能に大きな影響を与える。このため、通信装置内に実装されたTLBのヒット率向上が期待されている。

#### 【0014】

通信装置がアドレス変換を行なう従来例としては、例えば特許文献1ないし特許文献3が挙げられる。特許文献1（特開平05-089056号公報）には通信制御によりアドレス変換を行なう共有メモリにアクセスすることが開示されている。

特許文献2（特開平07-262146号公報）にはアドレス変換をルータで行い、パケットによりプロセッサ間通信を行なうことが開示されている。特許文献3（特許第3237599号公報）にはアドレス変換をプロセッサ内で行なうことが開示されている。

#### 【0015】

##### 【特許文献1】

特開平05-089056号公報

##### 【特許文献2】

特開平07-262146号公報

##### 【特許文献3】

特許第3237599号公報

#### 【0016】

##### 【発明が解決しようとする課題】

上述したように、複数の計算機を用いて並列処理を行なうためにTLB方式による通信装置を計算機に実装する場合には、TLBのヒット率の向上が重要となる。

**【0017】**

上述したように、通信装置内のTLBの実装には一般的には、変換対象論理アドレスと各TLBエントリ内の論理アドレスを比較するダイレクトマップ方式、セット・アソシアティブ方式、あるいはフル・アソシアティブ方式のいずれかが使用される。

**【0018】**

上記3つの方式では変換対象論理アドレスにより変換に使用するTLBエントリが決定される。また、TLBのエントリ数はアドレス変換テーブルのサイズよりも小さいため、TLBにミスし、かつ、TLBに空きエントリがない場合に、TLBエントリの置き換えが実施される。

プロセッサの使用する論理アドレスには局所性が高く、また、最近使用されたものの程、再度使用される可能性が高いという挙動があるため、置き換えアルゴリズムとしてはLRU (Least Recently Used) アルゴリズム、すなわち、最近、最も使われていないTLBを置き換えるアルゴリズムが使用されることが多い。

**【0019】**

ただし、データ転送においては、書き込み対象アドレスが単調に増加することが多く、過去に使用された主記憶アドレスが再利用される可能性が低いという、プロセッサのアクセスとは大きく異なる挙動を示すため、TLBエントリの置き換えアルゴリズムをLRUとすると、TLBヒット率が低下するという第一の問題がある。

。

**【0020】**

さらに、通信装置の受信部が保持するTLBは複数の通信相手に使用されるため、論理アドレスのみで使用されるTLBエントリを使用すると、一つの通信相手の送信装置からのデータ転送により通信装置の受信部のTLBエントリがすべて使用され、他の通信相手の送信装置からのデータ通信で使用していたTLBエントリが置き換えられるため、TLBヒット率が低下するという第二の問題がある。

**【0021】**

本発明は、複数の計算装置が通信装置を介して他の計算装置に通信する計算機システムにおいて、ユーザ・レベル通信を実現するためのアドレス変換装置としてTLBを通信装置内に実装した装置のTLBのヒット率を向上し、効率的に使用することのできる方式の提供を目的とする。

#### 【0022】

##### 【課題を解決するための手段】

本発明のデータ転送方法およびネットワークシステムは、複数の計算機がパケット交換網を介して接続されるネットワークシステムで行なわれるデータ転送方法において、

前記複数の計算機のそれぞれに、

主記憶装置と、

通信処理命令を発行するプロセッサと、

前記プロセッサからの通信命令を処理し、前記パケット交換網を介して他の計算機と通信を行う通信装置とを設け、

前記通信装置には、

交換網に通信を送信する送信部と交換網から通信を受信する受信部それぞれがTLBエントリを複数保持するTLBを設け、前記プロセッサは送り元の計算機の情報を含む前記通信命令を発行し、前記受信部は、使用するTLBエントリを送り元の計算機に応じて決定させることを特徴とする。

#### 【0023】

前記送り元計算機の情報は、送信部が通信命令に含まれる送り元計算機情報を通信パケットに付加するか、あるいは通信装置が保持する自計算機の情報を通信パケットに付加する。

#### 【0024】

また、通信装置は、送り元の計算機から同一宛先に対する同時に処理する通信命令の数を制限して処理を行ない、送信部が同一の宛先に対して同時に複数の通信命令を処理することにより、送信をオーバーラップさせ、かつ、高いTLBヒット率を実現することができる。

#### 【0025】

さらに、通信装置は、送り元の計算機における処理命令の識別番号を通信パケットに付与することとし、受信部は、前記処理命令の識別番号が同じである通信パケットについては同一のTLBエントリを使用することにより、受信部が使用するTLBエントリ選択論理を簡素化することができる。

#### 【0026】

##### 「作用」

本発明では、通信装置の受信部に実装されるTLBのヒット率向上を目的とし、通信装置の受信部のTLBにおいてTLBエントリを通信要求の送り元計算機毎に独立した1エントリあるいは複数の専用TLBエントリを保持し、同一の送り元からの通信の宛先論理アドレスがTLBにミスし、該送り元計算機専用確保されたTLBエントリに空きがない場合、前記送り元計算機専用TLBエントリのみを置き換え対象とする方式としている。

#### 【0027】

前記方式により、他の送り元からの通信による不要なTLBエントリ追い出しを回避でき、かつ、TLBの置き換え対象を容易に選択することを可能にする。

#### 【0028】

さらに、通信要求の送り元計算機において、同一の宛先計算機に対する通信命令の同時処理命令数を、規定された数の命令以下に制限し、受信側の通信装置に上記同時処理命令数以上のTLBエントリを実装し、かつ、前記通信命令毎にTLBエントリを専用割り当てることにより、複数の送り元計算機からの、それぞれ複数の通信命令により送信される通信パケットの受信時に、前記の通信で使用中のTLBエントリの追い出しが発生する可能性を低下させ、TLBヒット率を向上させる。

#### 【0029】

##### 【発明の実施の形態】

次に、本発明の実施例について図面を参照して説明する。

#### 【0030】

図1は本発明によるシステムの一実施例の構成を示すブロック図である。

#### 【0031】

図1に示されるシステムは、計算機100a～100zが交換網500を介して相互に接続されて通信を行うシステムであり、計算機100a～100zのそれぞれは、演算処理および通信処理の指示を行うプロセッサ200a～200z、プロセッサ200a～200zが処理を実施するためのデータを保持する主記憶装置300a～300zと、プロセッサ200a～200zからの通信命令を処理して交換網500を通して通信処理を行う通信装置400a～400zから構成され、通信装置400a～400zはそれぞれ論理アドレスを物理アドレスに変換するためのRTLB410a～410zと、STLB470a～470zを内包している。

#### 【0032】

図12は各計算機の内部構造を示すブロック図であり、図12では代表して計算機100a、100bが示され、内部構造について計算機100aのみが示されている。また、各計算機の動作はすべて同様のものであるため、以下の説明では、図示されているプロセッサ200a、主記憶装置300a、主記憶装置300a内のアドレス変換用のページテーブル310a、通信装置400a、RTLB410a、STLB470a、について、それぞれプロセッサ200、主記憶装置300、ページテーブル (Page Table) 310、通信装置400、RTLB410、STLB470として説明する。

#### 【0033】

通信装置400は、送信部450と、送信部用TLBであるSTLB470と、受信部用TLBであるRTLB410と、受信部460と、から構成されている。

#### 【0034】

送信部450は、プロセッサ200からの指示により、通信命令を主記憶装置300から取り出し、通信命令内の送り元データの論理アドレスを、ページテーブル310の一部の内容を保持するSTLB370により物理アドレスに変換した後、送信データを主記憶装置300から取り出し、前記取り出したデータと通信命令から通信パケットを生成し、交換網500を経由して宛先計算機に送付する。

**【 0 0 3 5 】**

S T L B 4 7 0 は、送信部 4 5 0 から受け取った論理アドレスを物理アドレスに変換するために前記論理アドレスと、主記憶装置 3 0 0 内のページテーブル 3 1 0 から取り出したページ属性を含むページエントリを一時的に保持する。

**【 0 0 3 6 】**

R T L B 4 1 0 は、通信パケット内の宛先論理アドレスを物理アドレスに変換するために宛先論理アドレスと、主記憶装置 3 0 0 内のページテーブル 3 1 0 から取り出したページ属性を含むページエントリを一時的に保持する。

**【 0 0 3 7 】**

受信部 4 6 0 は、交換網 5 0 0 から通信パケットを受け取り、R T L B 4 1 0 を使用して宛先論理アドレスを宛先物理アドレスへの変換を行った後、宛先物理アドレスに前記通信パケット内のデータを主記憶装置 3 0 0 に書き込む。

**【 0 0 3 8 】**

図 4 を用いてページテーブル 3 1 0 の内容を説明する。通信を実施するそれぞれのプロセス毎にページテーブルを確保し、それぞれを主記憶装置 3 0 0 上に配置し、ページテーブル 3 1 0 を論理アドレスから物理アドレスへの変換に使用する。論理アドレスはプロセス毎に独立に存在するため、プロセス番号により使用するページテーブル 3 1 0 を決定する。論理アドレスは予め決められたサイズのページ内のオフセットと、VPN(Virtual Page Number)と呼ぶ論理ページ番号に分割する。

例えば、64bit の論理アドレス(63-0)に対して 4KByte のページサイズである場合、VPN は論理アドレス(63-12)を示し、オフセットは論理アドレス(11-0)を示す。アドレス変換後の物理アドレスは、VPN を PPN(Physical Page Number)と呼ぶ物理ページ番号に置き換えたものであり、オフセットは変更されない。

**【 0 0 3 9 】**

ページテーブル 3 1 0 は複数のページ・エントリ 3 3 0 から構成され、前記 VPN が使用する T L B のページ・エントリ番号を示す。すなわち、VPN が 3 0 であれば、3 0 番の T L B のページ・エントリが使用される。ページ・エントリ 3 3 0 は、アドレス変換後のページ番号である PPN 3 2 0 と、アドレス変換後のページ

がリード可能であることを示すリード可能属性 322 と、前記ページがライト可能であることを示すライト可能属性 323 と、ページ・エントリの内容が有効であることを示すエントリ有効フラグ 321 と、から構成される。

#### 【0040】

図 10 は、プロセッサ 200 が通信装置 400 に指示する通信命令書式の一例を示す図である。

#### 【0041】

一つの通信命令 800 は、命令の機能コードを示す命令コード 810、宛先の計算機を示す宛先計算機番号 811、宛先のプロセスを示す宛先プロセス番号 812、宛先プロセスにおけるデータの書き込みアドレスを示す宛先論理アドレス 813、転送するデータの長さを示す転送長 814、送り元計算機番号 815、転送するデータの送り元プロセス内での論理アドレスを示す送り元アドレス 816、送り元のプロセスを示すプロセス番号 817 から構成される。

命令コード 810 には、宛先計算機内の論理アドレスを指定する場合は Remote DMA Write の機能コードを、宛先計算機内の論理アドレスを受信部が内部に保持しているアドレスを使用する場合は Send の機能コードを指定し、前記命令コード 810 が Send を示している場合は、前記宛先論理アドレス 813 は無効である。

#### 【0042】

図 11 は、図 10 に示された通信命令 800 を受けた通信装置 400 が交換網 500 に送信を行い、あるいは、交換網 500 から受信する通信パケットの書式の一例を示す図である。

#### 【0043】

図 10 に示された一つの通信命令 800 は、通信の信頼性の確保と通信装置 400 の構造簡素化のために、複数の通信パケット 900 に分割して送信する。

#### 【0044】

通信パケット 900 は、通信命令の命令コード 810 を設定した命令コード 910 と、通信命令 800 の宛先計算機番号 811 を設定する宛先計算機番号 911 と、通信命令 800 の宛先プロセス番号 812 を設定する通信相手のプロセスを示すプロセス番号 (PPID: Packet Process ID) 912 と、該パケットの宛先論理

アドレスを設定する宛先プロセス内のデータの書き込み先を示す宛先論理アドレス(PVA:Packet Virtual Address) 9 1 3 と、該パケットのデータの長さを設定するデータ長 9 1 4 と、通信命令 8 0 0 の送り元計算機番号 8 1 5 を設定する送り元識別のための送り元計算機番号 9 1 5 と、宛先に書き込まれるデータ 9 1 7 から構成される。

#### 【 0 0 4 5 】

通信装置 4 0 0 の交換網 5 0 0 への送信部 4 5 0 において、同時に同一宛先に対する通信命令最大同時処理数を 2 個以上とする場合は、送信部 4 5 0 は、同時に処理する命令の識別番号を示す命令識別番号 9 1 6 を通信パケットに追加し、受信部 4 6 0 において使用する T L B エントリを指定することにより、送信部 4 5 0 が同一宛先に対する複数通信命令の同時処理時の T L B ヒット率を向上させる。

#### 【 0 0 4 6 】

図 2 は、通信装置 4 0 0 が実装する R T L B 4 1 0 の構成を示す図である。R T L B 4 1 0 は交換網からの受信パケットの送り元計算機毎にそれぞれ 1 エントリあるいは複数エントリが用意された複数の T L B エントリから構成されている。1 つの T L B エントリは T L B エントリ有効フラグ 4 2 0、プロセス番号 (TP ID:TLB Process ID) 4 2 1、仮想ページ番号 (TVPN:TLB Virtual Page Number) 4 2 2、および物理ページ番号 (TPPN:TLB Physical Page Number) 4 2 3 から構成されている。

#### 【 0 0 4 7 】

図 3 は、通信装置 4 0 0 における論理アドレスを物理アドレスに変換するアドレス変換方式を示す図である。

#### 【 0 0 4 8 】

図 1 1 に示した通信パケット 9 0 0 を受け付けた受信部 4 6 0 は、宛先プロセス番号 (PPID:Packet Process ID) 9 1 2 と、宛先論理アドレス (PVA:Packet Virtual Address) 9 1 3 とから物理アドレスを生成する。

#### 【 0 0 4 9 】

PVA 9 1 3 は仮想ページ番号 (PVPN:Packet Virtual Page Number) 9 2 1 と



、オフセット (Offset) 9 2 2 とから構成される。PPID 9 1 2 およびPVPN 9 2 1 が、通信パケット 9 0 0 に対応する T L B エントリの TPID 4 2 1 および TVPN 4 2 2 と比較され、上記比較結果が共に一致し、かつ、該 T L B エントリ有効フラグ 4 2 0 が有効を示していた場合に、T L B ヒットとし、TPPN 4 2 3 と Offset 9 2 2 から変換後物理アドレス 9 3 2 が生成される。

#### 【0050】

上記手順によりPPID 9 1 2 と、PVA 9 1 3 とから物理アドレス 9 3 2 にアドレス変換が行われる。

#### 【0051】

図 4 は、上記比較結果が不一致であった場合、あるいは T L B エントリ有効フラグ 4 2 0 が有効を示していなかった場合、すなわち T L B ミスの場合のアドレス変換手順を説明するための図である。

#### 【0052】

通信装置 4 0 0 を使用するソフトウェアは、論理アドレスを物理アドレスに変換するためのページテーブル 3 1 0 を主記憶装置 3 0 0 内の通信装置 4 0 0 がアクセス可能な領域内に予め確保する。1 プロセスに対し 1 つのページテーブルが使用される。ページテーブル 3 1 0 は、複数のページ・エントリ 3 3 0 から構成され、前記ページ・エントリの 1 エントリは 1 ページのアドレス変換に使用され、変換する論理アドレスの一部である受信パケットの P V P N 9 2 1 の値毎に 1 つの物理ページ番号 (PPN: Physical Page Number) 3 2 0 と、アドレス変換後のページがリード可能であることを示すリード可能属性 3 2 2 と、前記ページがライト可能であることを示すライト可能属性 3 2 3 と、ページ・エントリ 3 3 0 の内容が有効であることを示すエントリ有効ビット 3 2 1 と、から構成される。

#### 【0053】

アドレス変換によって生成される物理アドレスは前記 P P N 3 2 0 と、論理アドレスのページ内アドレスであるオフセットを組み合わせることによって生成される。例えば、論理アドレスが 6 4 ビット、ページサイズが 4 KByte、物理アドレスが 5 0 ビットであった場合、前記 P P N 3 2 0 が物理アドレス (49-12) となり、前記オフセット (11-0) が物理アドレス (11-0) となる。ここで ( ) 内はビット位置

を示し()内の“-”はビット範囲を示し、“-”の左側がMSB(Most Significant Bit), “-”の右側がLSB(Least Significant bit)を示す。つまり(63-12)はビット位置63からビット位置12までの52ビットを示す。

#### 【0 0 5 4】

前記TLBミスが発生した場合、通信装置400は通信パケット900のPPID912、PVPN921をTLBミスしたエントリのそれぞれTPID421、TVPN422に登録し、PVPN921によりページ・テーブル310から読み出したページ・エントリ330を上記TLBエントリのTPPN423更新し、TLB有効フラグ420に前記読み出したページ・エントリ有効フラグ321の値をセットする。

#### 【0 0 5 5】

上記TLBエントリ有効フラグ420をセットした後、通信パケット900により、RTL410を再索引する。

#### 【0 0 5 6】

TPID421、TVPN422には通信パケット900内のそれぞれPPID912、PVPN921が登録されており、TLBエントリ有効フラグ420が有効を示す場合は、通信パケット900はTLBにヒットする。再索引によるTLBヒット後のアドレス変換は初回索引時にTLBにヒットした場合のアドレス変換と同様である。前記TLBエントリ有効フラグ420が無効を示す場合はアドレス変換の例外が発生し、パケットを破棄する、あるいはパケット送信元に不正を通知する等の例外処理を実施する。

#### 【0 0 5 7】

図5は、TLB索引アドレスを計算する方法を説明するための図である。

#### 【0 0 5 8】

受信部460が、送り元計算機番号915、PPID912、PVA913および宛先に書き込むための通信データ(Data)917を含む通信パケット900を受け取った後、送り元計算機番号915の値がNであったとすると、TLBエントリNを取り出す。

#### 【0 0 5 9】

PVA913はPVPN921とOffset922から構成される。TLBエントリNは

送り元計算機専用に確保されたTLBエントリであり、他の送り元計算機からの通信によって追い出されることはない。宛先アドレスが単調増加することが多いデータ転送の特徴を考慮すると、上記TLBエントリは、宛先アドレスのPVPN 921の値に変更があるまではTLBエントリNにヒットし続ける。前記のTLB構成により、他の送り元計算機からのアクセスによって有効なTLBエントリが追い出されない。

#### 【0060】

図6は、命令識別番号916を用いて、通信装置400の送信部450が同一宛先に対して複数の通信命令を処理する実施例を説明するための図である。

#### 【0061】

本実施例に示される通信パケット900'は、図5で説明したTLB索引方式で用いられた通信パケット900に、送信部が同一宛先に対して同時に処理する通信命令の識別番号である916が追加し、TLBの索引アドレスの生成に使用することを特徴とし、TLB索引アドレスNに、通信パケット900'内の命令識別番号916を使用すること以外は上記図5で説明した方式と同様である。

#### 【0062】

通信パケット900'を受け取った通信装置400の受信部460は、送り元計算機番号915の値がj、命令識別番号916の値がk、通信装置400の送信部が同一宛先に対して同時に処理する通信命令の数をPmaxとすると、TLBエントリの索引アドレスNを下記の式(1)により計算する。

#### 【0063】

$$N = P_{\max} \times j + k \quad (1)$$

TLB索引アドレスNに通信パケット900'の命令識別番号916を使用することにより、通信装置400の送信部450からの同一宛先計算機の通信装置400へ同時に複数命令により生成された通信パケットを入れ換えて送信した場合にもTLBエントリの追い出しが避けられ、TLBのヒット率を向上させることができる。

#### 【0064】

図13を用いてTLB索引および置き換え動作を説明する。

## 【0065】

S100：TLB索引開始。交換網500から受信したパケット900の論理アドレス変換動作を開始する。

## 【0066】

S110："TLB索引アドレス生成"によりTLBの索引アドレスを生成する。前記同時処置通信命令数をPmax、通信命令識別番号916の値をk、送り元計算機番号をjとすると、前記式(1)から、

$$\text{TLB索引アドレス} = \text{Pmax} \times j + k$$

となる。前記TLB索引アドレスに登録されているTLBエントリを対象エントリとして以下の処理を実施する。

## 【0067】

S120："PVPN = TVPN?"において、TLBのヒット判定が実施され、前記受信パケット900の宛先プロセス番号PPID912と、宛先論理アドレス913の仮想ページ番号であるPVPN921とを、それぞれTLBの対象エントリのプロセス番号TPID421、仮想ページ番号TVPN422の値と比較する。前記PIDとVPNの両方が一致し、かつ、TLBエントリ有効フラグ420が有効を示していた場合はTLBヒットとなりS130へ、前記以外の場合はTLBミスとなりS200へ移行する。

## 【0068】

前記はTLBをダイレクトマップ方式で構成した場合の処理を示したものである。TLBをセット・アソシアティブ方式で構成した場合は、前記比較は全セットのTLBエントリのTPIDとTVPNとを、受信パケットのPVPNとPPIDと比較する。前記複数のTLBエントリ中、いずれか1つのTLBエントリの一致が検出され、かつ、該TLBエントリのTLBエントリ有効フラグ420の値が有効を示していた場合にTLBヒットとし、TLBヒットでない場合をTLBミスとする。

## 【0069】

さらに、TLBをフル・アソシアティブ方式で構成した場合は、全TLBエントリのTVPNと受信パケットのPVPNを比較すること以外は前記セット・アソシアティブ方式と同様である。

**【 0 0 7 0 】**

S 1 3 0：前ステップ S 1 2 0 にて T L B ヒットの場合は、T L B の対象エントリの物理ページ番号 T P P N 4 2 3 と、受信パケットの論理アドレスのページ内アドレスであるオフセット 9 2 2 の値から物理アドレスを生成する。

**【 0 0 7 1 】**

S 5 0 0：受信パケットの物理アドレス変換がステップ S 1 3 0 にて完了し、T L B 動作が終了する。

**【 0 0 7 2 】**

S 2 0 0：ステップ S 1 2 0 において T L B ミスした場合、T L B の対象エントリ有効フラグ 4 2 0 が有効を示しているかをチェックする。有効であった場合は S 2 1 0 へ、無効であった場合は S 2 2 0 へ移行する。

**【 0 0 7 3 】**

S 2 1 0：T L B 内の対象エントリのエントリ有効フラグ 4 2 0 を無効を示す値にクリアし、エントリを無効にする。

**【 0 0 7 4 】**

前記は T L B をダイレクトマップ方式で構成した場合を示す。T L B をセット・アソシアティブ方式あるいはフル・アソシアティブ方式で構成した場合は、複数の T L B エントリ中の 1 エントリを L R U (Least Recently Used) 方式等で選択し、前記選択されたエントリ有効フラグ 4 2 0 を無効を示す値にクリアし、以後の処理対象エントリとする。

**【 0 0 7 5 】**

S 2 2 0：T L B 内の対象エントリの論理ページ番号 T V P N 4 2 2 に受信パケットの P V P N 9 2 1 の値を設定する。

**【 0 0 7 6 】**

S 2 3 0：ページ・テーブル 3 1 0 からページ・エントリ 3 3 0 を取り出す。ページ・エントリ 3 3 0 は受信パケットの論理ページ番号 P V P N 9 2 1 により索引される。

**【 0 0 7 7 】**

S 2 4 0：ページ・エントリ 3 3 0 から取り出した、物理ページ番号 P P N 3

20および属性フラグのリード可能属性322と、ライト可能属性323と、ページ・エントリ有効フラグ321をTLBの対象エントリに登録する。

#### 【0078】

S250:前記読み出したページ・エントリ有効フラグ321の値が有効を示し、かつ、前記ライト可能属性323が”1”であった場合、再度TLBを索引するためS110へ移行する。前記エントリ有効フラグ321の値が無効を示すか、あるいは、前記ライト可能属性323が”0”、すなわちライト不可であった場合は、ページ変換が失敗し、受信パケットを破棄し処理を終了するため、S500へ移行する。

#### 【0079】

次に、通信装置400の同一宛先計算機への複数通信命令の同時並行処理について説明する。

#### 【0080】

図7に示されるように、通信装置400aが交換網500を通して通信装置400bに通信を行う。図8は、通信装置400aが処理しようとする通信命令を示しており、通信装置400aはプロセッサからの要求により、宛先が通信装置400bである通信命令1～4の処理を実行する。通信命令1、2、3、4はそれぞれ5、3、3、1個のパケットに分割されて交換網500を通して通信装置400bに送られる。

#### 【0081】

図14は通信装置400aが交換網500を通して通信装置400bに送信し、複数の通信命令から生成される通信パケットが通信命令の順序で送信される場合のパケット送信例を示す。通信命令1、2、3、4はそれぞれ5個、3個、3個、1個のパケットを送信する。

#### 【0082】

送信順序1～5において通信命令1が通信パケット1-1～1-5の5パケットに分割して送信され、送信順序6～8において通信命令2が通信パケット2-1～2-3の3パケット、送信順序9～11において通信命令3が通信パケット3-1～3-3の3パケット、送信順序12において通信命令4が通信パケッ

ト 4 - 1 の 1 パケットが送信されている。通信命令 1、2、3、4 の最後のパケットはそれぞれ送信順序 5、8、11、12 に送信している。図 14 の例においては、同時通信命令処理数が”2”であり、通信パケット内命令識別番号が 1 あるいは 2 のいずれかの値を取り、通信命令 1、3 においては前記識別番号の値が 1、通信命令 2、4 においては前記識別番号の値が 2 である。

#### 【0083】

図 9 は、通信装置 400 a が交換網 500 を通して通信装置 400 b に送信し、複数の通信命令から生成される通信パケットがお互いに入り組んだ順序で送信された送信例を示す。

#### 【0084】

図 9 に示される例では、通信装置 400 a が同時に同一宛先に対して処理可能な通信命令の数が 2 であり、交換網 500 に送信するパケットの生成元の未完了通信命令数は最大”2”である。前記未完了通信命令とは、通信命令の最終パケットが交換網に送信されていない通信命令を意味する。また、図 9 の通信命令 1、2、3、4 はそれぞれ図 14 の通信命令 1、2、3、4 と同じ数のパケットに分割される。すなわち、通信命令 1、2、3、4 はそれぞれ 5 個、3 個、3 個、1 個のパケットを送信する。

#### 【0085】

送信順序 1 および 2 において、通信装置 400 a は通信命令 1 により生成された通信パケット 1 - 1、1 - 2 についての命令識別番号を”1”として、交換網 500 に送り、命令識別番号”1”を使用中とする。

#### 【0086】

送信順序 3、4 において、通信装置 400 a は通信命令 1 の全パケットを未送出であるが、通信命令 2 から生成された通信パケット 2 - 1、2 - 2 の準備ができており、かつ、通信命令 2 を処理しても同時送信命令数制限以下であるため、通信パケット 2 - 1、2 - 2 を既に使用中の命令識別番号である”1”とは異なる命令識別番号、すなわち命令識別番号”2”を付加して、交換網 500 に送付し、命令識別番号”2”を使用中とする。

#### 【0087】

送信順序 5 では、通信装置 400 a は通信命令 1 用の通信パケット 1-3 を送付する。

【0088】

送信順序 6 において、通信装置 400 a は通信命令 2 から生成された最後の通信パケット 2-3 を交換網 500 に送付する。通信パケット 2-3 は通信命令 2 から生成される最後のパケットであるため、通信パケット 2-3 送付後に通信装置 400 a の宛先 400 b に対する同時通信命令処理数が 1 個となり、宛先 400 b に対して新たに通信命令の処理が実行可能となり、命令識別番号”2”を未使用状態とする。

【0089】

送信順序 7 において、通信装置 400 a は新たに通信命令 3 の処理を開始し、通信パケット 3-1 に未使用の命令識別番号である命令識別番号”2”を付加して交換網 500 に送付する。送信順序 8, 9 においては通信命令 1 から生成された通信パケット 1-4, 1-5 を送付する。通信パケット 1-5 は通信命令 1 から生成された最後の通信パケットであるため、通信命令 1 のパケット送信処理が完了し、通信装置 400 a の宛先通信装置 400 b に対する同時命令処理数が 1 個になり、命令識別番号”1”が未使用となる。

【0090】

送信順序 10 において、通信装置 400 a は通信命令 3 の通信パケット 3-2 が交換網 500 に送付する。

【0091】

送信順序 11 において、通信装置 400 a は新規に通信命令 4 の処理により生成された通信パケット 4-1 が未使用命令識別番号”1”が付加されて交換網 500 に送付される。通信パケット 4-1 は通信命令 4 の最後の通信パケットであるため、命令識別番号”1”は未使用状態となる。

【0092】

送信順序 12 において、通信命令 3 の最後の通信パケット 3-3 を交換網 500 に送付し、命令識別番号”2”を未使用状態とする。

【0093】



次に、上記 T L B が追い出されることによる T L B ヒット率低下の問題を解決する通信パケット送信部 4 5 0 における処理について説明する。

#### 【 0 0 9 4 】

一つの通信命令においては、宛先論理アドレスは単調に増加することが多いというデータ転送の特徴、および通信パケットの受信部において、T L B エントリが送り元毎に決められた個数だけ確保されていることを考慮し、送信部 4 5 0 では同一宛先に対する同時に処理する通信命令の個数を上記制限数以下に抑えることにより、上記 T L B ヒット率を向上させる。さらに、送信部 4 5 0 が同時に処理する命令が使用する受信部 4 6 0 の T L B エントリを指定するため、各パケットに命令識別番号 9 1 6 を付加することにより、同一通信命令 8 0 0 により生成された通信パケット 9 0 0 が通信装置 4 0 0 の受信部 4 6 0 の同一 T L B を使用する方式とする。

#### 【 0 0 9 5 】

図 1 7 を用いて送信部 4 5 0 の構成を詳細に説明する。送信部 4 5 0 はプロセッサから送信命令処理開始指示を受け取り、主記憶装置 3 0 0 から送信命令を取り出し、前記取り出した送信命令をパケット生成部に送信する命令処理部 4 5 1 と、命令処理部から送信命令を受け取り、送信データバッファの論理アドレスを S T L B 4 7 0 を使って物理アドレスに変換した後に主記憶装置 3 0 0 から送信データを読み出し、前記読み出したデータから送信パケットを生成し、交換網 5 0 0 に送信するパケット生成部 4 5 2 から構成される。

#### 【 0 0 9 6 】

図 1 8 を用いて同時送信命令数制限を 2 とした場合の送信部の動作を詳細に説明する。

#### 【 0 0 9 7 】

命令処理開始：処理中命令数の初期値は” 0 ”である。

#### 【 0 0 9 8 】

N 0 1 0 ：プロセッサ 1 0 0 から命令処理部 4 5 1 に通信命令処理指示があった場合は N 1 0 0 へ。

前記以外の場合は N 0 1 0 へ。

**【 0 0 9 9 】**

N 1 0 0 : 命令処理部 4 5 1 が主記憶装置 3 0 0 から送信命令取り出す。

**【 0 1 0 0 】**

N 1 1 0 : 命令処理部 4 5 1 が既にパケット生成部 4 5 2 へ送信指示をした送信命令の中で、新たに送信を完了した送信命令があれば N 1 2 0 へ。前記完了済み送信命令がなければ N 1 3 0 へ。

**【 0 1 0 1 】**

N 1 2 0 : パケット生成部 4 5 2 から送信命令の完了通知を受け取り、前記完了通知の数を命令処理部 4 5 1 内の処理中命令数から減算する。処理中命令数 - 完了命令数は、処理中命令数 = 処理中命令数 - 完了命令数を省略表記したもの。

**【 0 1 0 2 】**

N 1 3 0 : 処理中命令数が 2 未満であれば N 1 4 0 へ。前記以外の場合は N 1 1 0 へ。

**【 0 1 0 3 】**

N 1 4 0 : 命令処理部 4 5 1 からパケット生成部 4 5 2 へ通信命令を転送する。

**【 0 1 0 4 】**

命令処理終了 : 1 つの命令に対する命令処理部 4 5 1 の処理終了。

**【 0 1 0 5 】**

命令処理部 4 5 1 から通信命令を受け取ったパケット生成部 4 5 2 は、通信命令 8 0 0 内の送り元論理アドレス 8 1 3 を S T L B 3 7 0 を使用して物理アドレスに変換し、転送長 8 1 4 の長さの送信データを主記憶装置 3 0 0 から取り出し、前記データを規定の長さに分割し、分割されたそれぞれのデータからパケットを生成して交換網 5 0 0 へ送信する。

**【 0 1 0 6 】**

同時命令発行数制限はハードウェアで実装する T L B のエントリ数が最も大きな制約となる。すなわち、送信部が同時に処理可能な命令数は受信側および送信側に実装可能な T L B のエントリ数に依存する。送信部は同時処理可能な命令数

の TLB エントリを保持すれば十分な TLB のヒット率が得られる。受信側は、同時処理可能な命令数×送り元計算機の数 of TLB エントリ、を保持すれば十分な TLB のヒット率が得られる。

例えば、同時処理命令数が 2、交換網 500 で 10 台の計算機を接続したとすると、受信部の TLB エントリ数としては  $2 \times 10 = 20$  エントリ実装すれば十分である。また、同時処理命令数が 1 では大きなデータサイズの転送を実施した場合に、後続の小さいデータサイズの転送が長時間待たされることになるため、通常は同時処理命令数を 2 以上とする。

#### 【0107】

次に、図 15、図 16 を用いて同時命令数制限によりヒット率が向上する効果について説明する。

#### 【0108】

先に図 15 を用いて、同時送信命令制限のない場合の受信側 TLB の動作を詳細に示す。

#### 【0109】

図 15 では受信側の TLB エントリ数を 1、送信命令 1 と送信命令 2 がそれぞれ 2 個のパケットを送信し、送信命令 1 がパケット 1-1 とパケット 1-2、送信命令 2 がパケット 2-1 とパケット 2-2 を生成した場合の動作を示す。送信命令 1 と送信命令 2 の宛先アドレス 813 は異なるページを使用し、パケット 1-1 とパケット 1-2 は同一ページ内のアドレスへの転送とし、パケット 2-1 とパケット 2-2 は同一ページ内のアドレスへの転送とする。

#### 【0110】

開始

T11：パケット受信前の TLB は無効状態。

P11：パケット 1-1 を受信

T12：TLB ミス（ミス 1 回目）

T13：TLB に送信命令 1 用の変換情報を登録する。

P12：パケット 2-1 を受信

T14：TLB ミス（ミス 2 回目）。送信命令 2 と送信命令 1 は異ページを使用

する。

T15: TLBには送信命令2用の変換情報を登録する。

P13: パケット1-2を受信

T16: TLBミス(ミス3回目)。送信命令1と送信命令2は異ページを使用する。

T17: TLBに送信命令1用の変換情報が登録される。

P14: パケット2-2を受信。

T18: TLBミス(ミス4回目)。送信命令2と送信命令1は異ページを使用する。

T19: TLBには送信命令2用の変換情報を登録する。

終了

本動作においてTLBミスが4回発生した。

#### 【0111】

図16を用いて、同時送信命令制限のある場合の受信側TLBの動作を詳細に示す。

#### 【0112】

図16では送信側の同時命令制限数が1であること以外は図15と同じであり、受信側の

TLBエントリ数1、送信命令1と送信命令2がそれぞれ2個のパケットを送信し、送信命令1がパケット1-1とパケット1-2、送信命令2がパケット2-1とパケット2-2を生成した場合の動作を示す。送信命令1と送信命令2の宛先アドレスは異なるページを使用し、パケット1-1とパケット1-2は同一ページ内のアドレスへの転送とし、パケット2-1とパケット2-2は同一ページ内のアドレスへの転送とする。

#### 【0113】

開始

T21: パケット受信前のTLBは無効状態。

P21: パケット1-1を受信

T22: TLBミス(ミス1回目)

T23：TLBに送信命令1用の変換情報を登録する。

P22：パケット1-2を受信

T24：TLBヒット（ヒット1回目）。パケット1-1とパケット1-2は同ページ使用。

T25：TLBには送信命令1用の変換情報が登録済み。

P23：パケット2-1を受信

T26：TLBミス（ミス2回目）。送信命令1と送信命令2は異ページを使用する。

T27：TLBに送信命令2用の変換情報が登録される。

P24：パケット2-2を受信。

T28：TLBヒット（ヒット2回目）。パケット2-1とパケット2-2は同ページ使用。

T29：TLBには送信命令2用の変換情報が登録済み。

終了

本動作において、TLBミス2回と、TLBヒット2回が発生した。

#### 【0114】

命令数制限をしない図15ではTLBヒットが4パケット中0回、すなわちTLBヒット率は0%。一方、命令数制限をした図16においてはTLBヒットが4パケット中2回、すなわちTLBヒット率50%。従って、命令数制限によりヒット率を向上させる効果が見られた。

#### 【0115】

##### 【発明の効果】

本発明は上記のように構成されているので、以下に記載するような効果を奏する。

#### 【0116】

複数の送り元からの通信パケットに対して、それぞれ独立にTLBのエントリを割り当てることにより、異なる送り元通信装置からの通信パケットの宛先論理アドレスによってTLBエントリが追い出されない効果がある。

#### 【0117】

また、通信パケットの送り元計算機からの同一宛先計算機へ同時に処理する通信命令の数を制限する場合には、通信パケットの受信側においては、一つの送り元に対して保持すべき T L B エントリの数を上記制限以下に押さえることが可能となるとともに、上記同時処理コマンド数分の T L B エントリを保持される。これにより、常に送り元計算機が送信中の通信命令に対応する T L B エントリを保持し続けることとなり、T L B ヒット率を向上させる効果がある。

#### 【0118】

また、通信パケット上に送り元通信装置の処理命令識別番号を付加し、受信部で使用する T L B エントリを指定する場合には、送り元通信装置が処理中の 1 つの通信命令が常に同一 T L B エントリを使用することに加えて、データ転送における 1 つの通信命令の宛先論理アドレスが単調増加する可能性が高いことから、送り元が同時に複数の同一宛先への通信命令の処理を実行させることが可能となり、かつ、T L B ヒット率を向上させる効果がある。

#### 【図面の簡単な説明】

##### 【図 1】

本発明の計算機システム

##### 【図 2】

T L B 構成

##### 【図 3】

アドレス変換方式

##### 【図 4】

ページテーブル構成

##### 【図 5】

T L B 索引方式その 1

##### 【図 6】

T L B 索引方式その 2

##### 【図 7】

通信方式

##### 【図 8】

## 通信装置の処理命令例

### 【図 9】

## 通信装置の packets 送信例

### 【図 1 0】

## 通信命令書式

### 【図 1 1】

## 通信 packets 書式

### 【図 1 2】

計算機の内部構造を示すブロック図である。

### 【図 1 3】

ページテーブルの内容を説明するための図である。

### 【図 1 4】

ページテーブルの内容を説明するための図である。

### 【図 1 5】

同時命令数制限によりヒット率が向上する効果について説明するための図である。

### 【図 1 6】

同時命令数制限によりヒット率が向上する効果について説明するための図である。

### 【図 1 7】

送信部 4 5 0 の構成を詳細に示すブロック図である。

### 【図 1 8】

同時送信命令数制限を 2 とした場合の送信部の動作を示すフローチャートである。

### 【符号の説明】

- 1 0 0 a、1 0 0 b、1 0 0 z 計算機
- 2 0 0 a、2 0 0 b、2 0 0 z プロセッサ
- 3 0 0 a、3 0 0 b、3 0 0 z 主記憶装置
- 3 1 0 ページ・テーブル

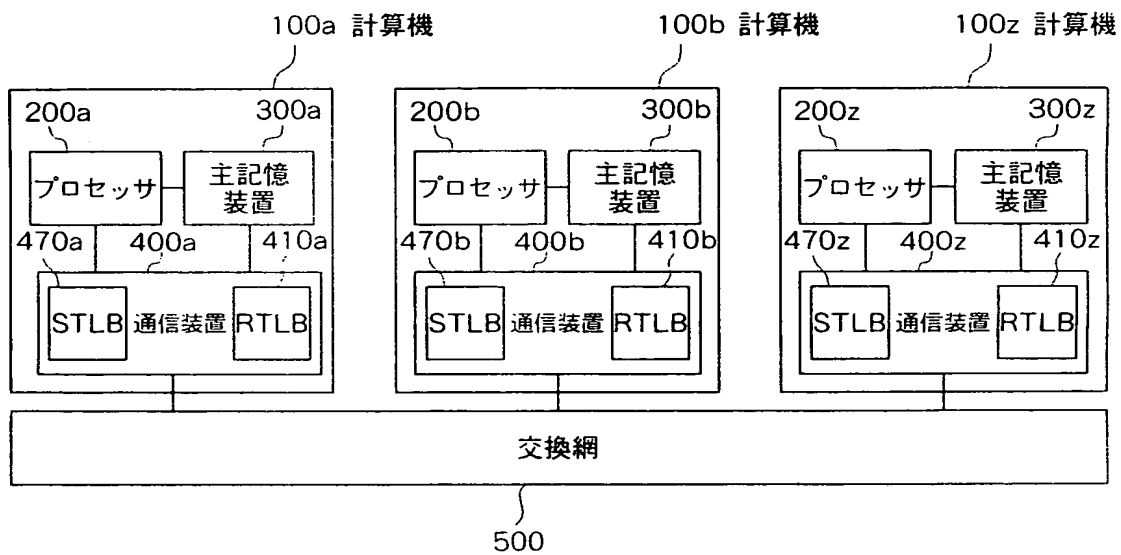
- 3 2 0 物理ページ番号
- 3 2 1 ページ・エントリ有効フラグ
- 3 2 2 ページ・リード可能属性
- 3 2 3 ページ・ライト可能属性
- 3 3 0 ページ・エントリ
- 4 0 0 通信装置
- 4 0 0 a、4 0 0 b、4 0 0 z 通信装置
- 4 1 0 a、4 1 0 b、4 1 0 z T L B (Translation Look aside Buffer)
- 4 1 0 T L B (Translation Look aside Buffer)
- 4 2 0 T L B エントリ有効フラグ
- 4 2 1 T P I D (T L B 内プロセスID)
- 4 2 2 T V P N (T L B 内仮想アドレスページ番号) ( T L B Virtual Page Number)
- 4 2 3 T P P N (T L B 内物理ページ番号) (T L B Physical Page Number )
- 4 5 0 通信装置の送信部
- 4 5 1 送信部の命令処理部
- 4 5 2 送信部のパケット生成部
- 4 6 0 通信装置の受信部
- 4 7 0 送信部が使用する T L B
- 5 0 0 交換網
- 8 0 0 通信命令
- 8 1 0 命令コード
- 8 1 1 宛先計算機番号
- 8 1 2 宛先プロセス番号
- 8 1 3 宛先論理アドレス
- 8 1 4 転送長
- 8 1 5 送り元計算機番号
- 8 1 6 送り元プロセス番号



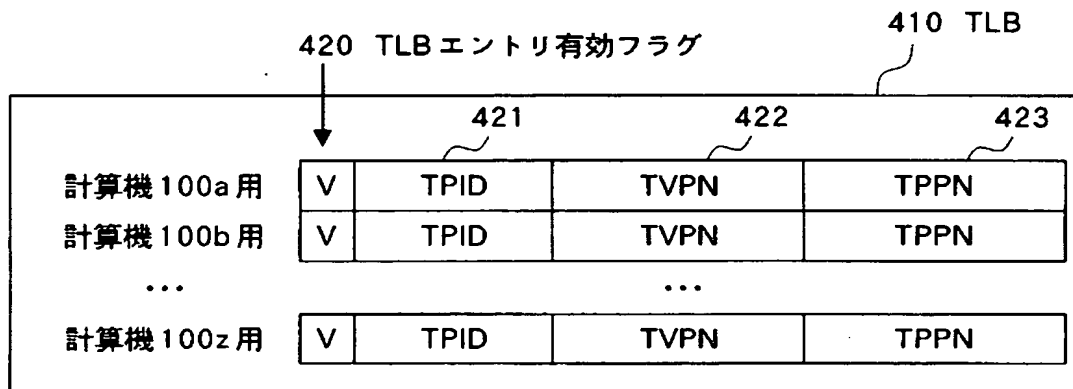
- 8 1 7 送り元論理アドレス
- 9 0 0 通信パケット
- 9 1 0 命令コード
- 9 1 1 宛先計算機番号
- 9 1 2 PPID(パケット・プロセス・ID)(Packet Process ID)
- 9 1 3 PVA(パケット仮想アドレス)(Packet Virtual Address)
- 9 1 4 パケット内データ長
- 9 1 5 送り元計算機番号
- 9 1 6 命令識別番号
- 9 1 7 転送データ
- 9 2 1 PVPN(パケット仮想ページ番号)(Packet Virtual Page Number)
- 9 2 2 Offset(仮想アドレス・オフセット)
- 9.3 2 変換後物理アドレス

【書類名】 図面

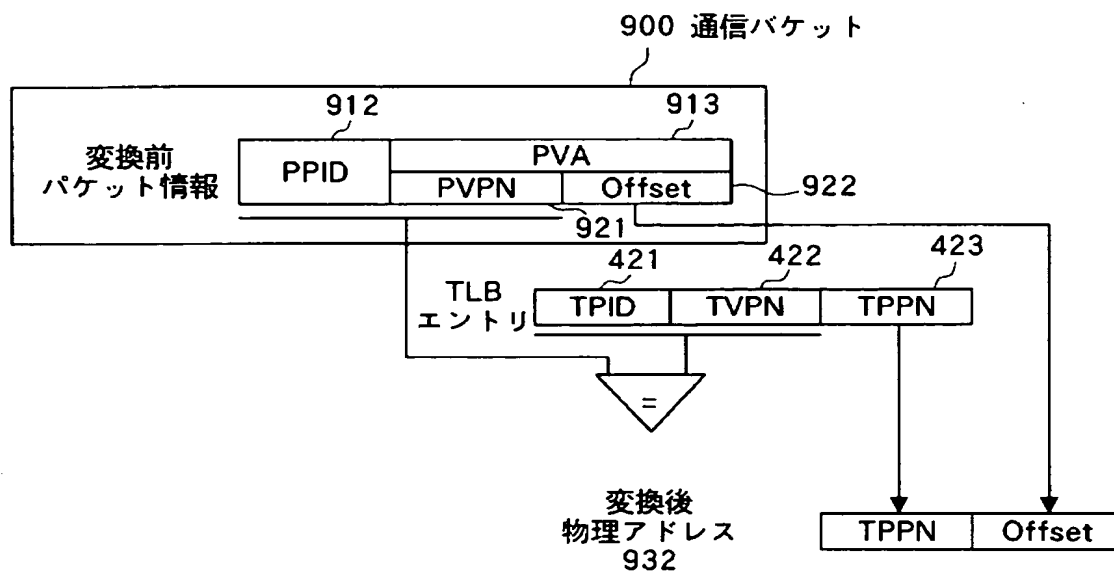
【図 1】



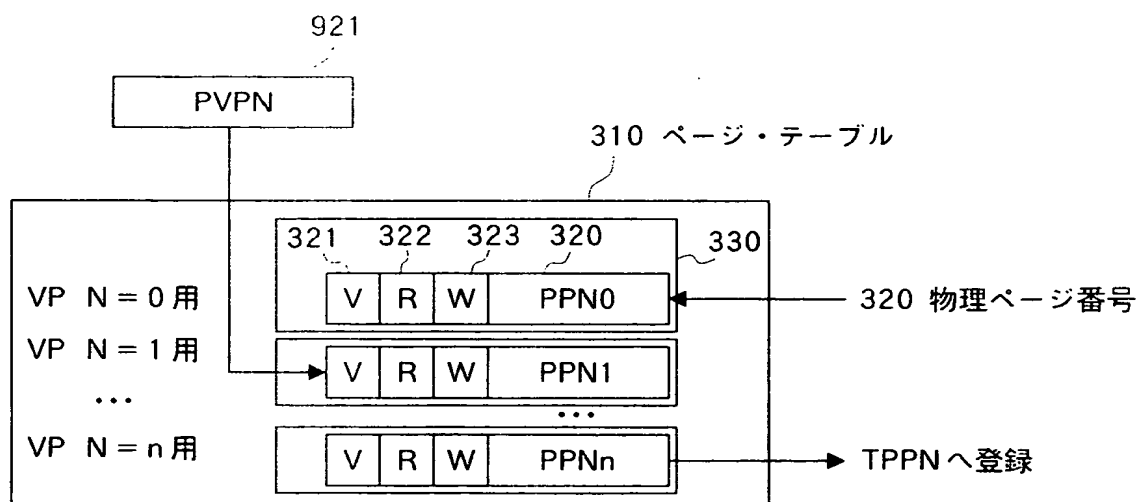
【図 2】



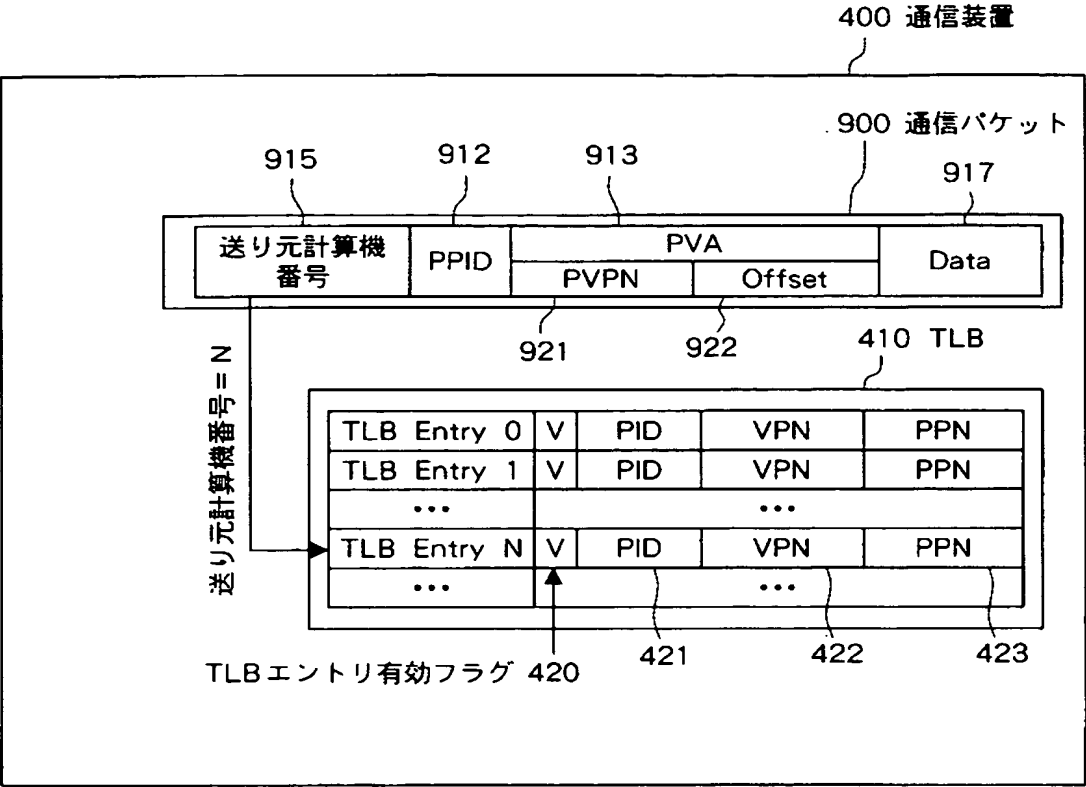
【図 3】



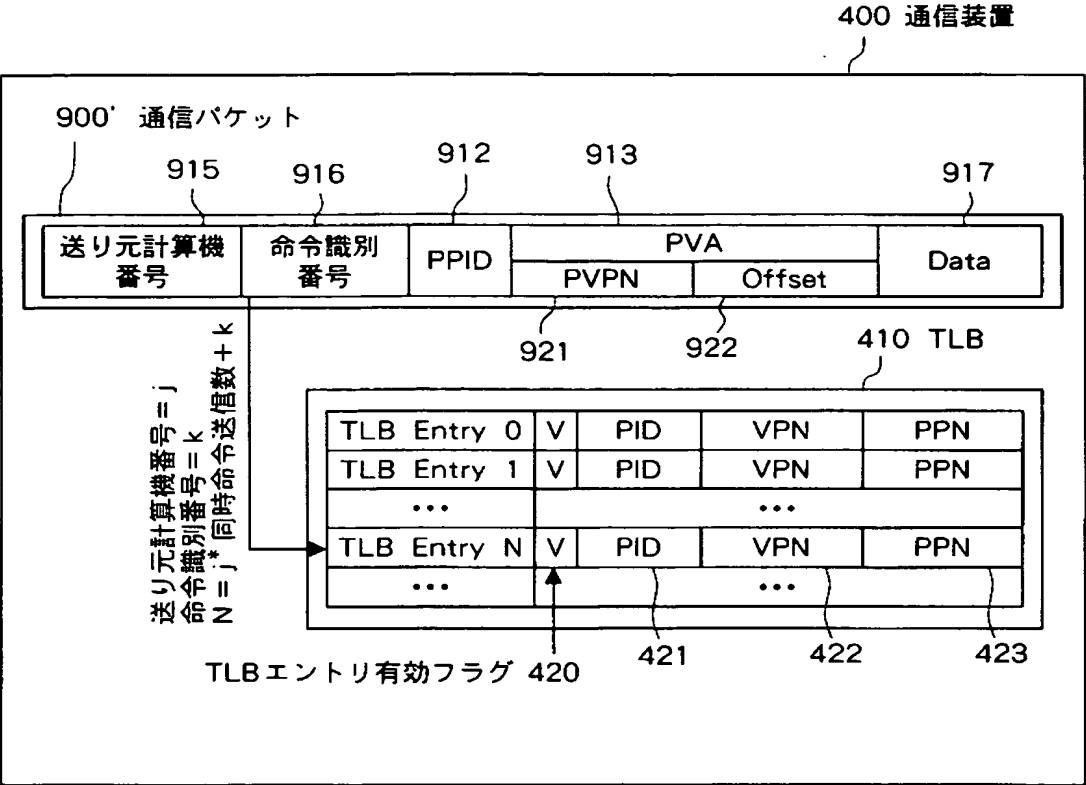
【図 4】



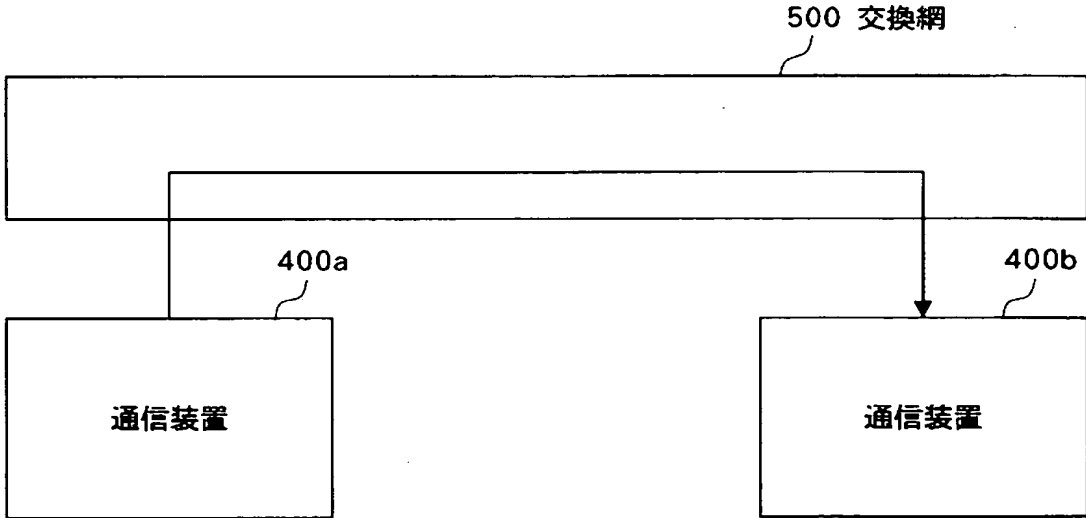
【図 5】



【図6】



【図7】



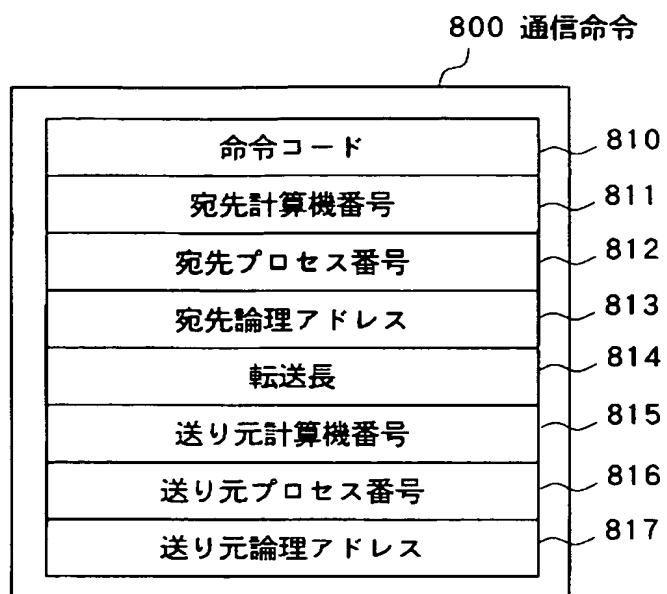
【図 8】

| 命令     | 宛先   | パケット数 |
|--------|------|-------|
| 通信命令 1 | 400b | 5     |
| 通信命令 2 | 400b | 3     |
| 通信命令 3 | 400b | 3     |
| 通信命令 4 | 400b | 1     |

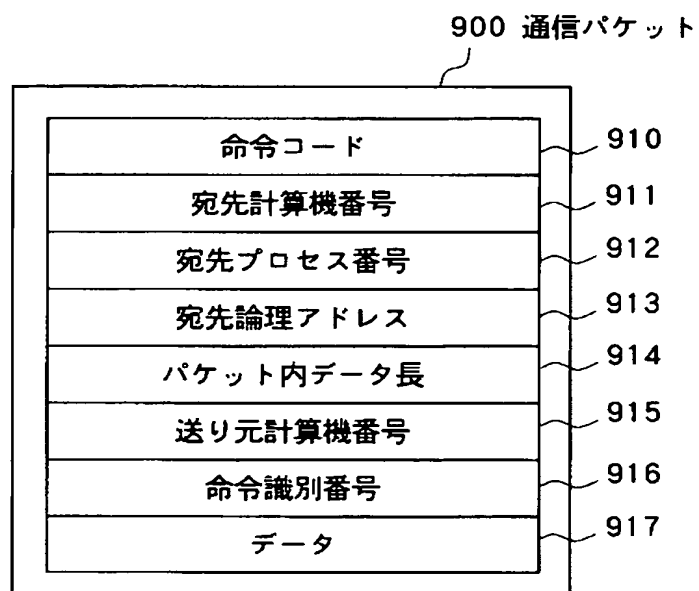
【図 9】

| 送信順序 | 通信命令   | 通信パケット     | 通信パケット内<br>命令識別番号 | 通信命令内<br>最終パケット |
|------|--------|------------|-------------------|-----------------|
| 1    | 通信命令 1 | 通信パケット 1-1 | 1                 | No              |
| 2    | 通信命令 1 | 通信パケット 1-2 | 1                 | No              |
| 3    | 通信命令 2 | 通信パケット 2-1 | 2                 | No              |
| 4    | 通信命令 2 | 通信パケット 2-2 | 2                 | No              |
| 5    | 通信命令 1 | 通信パケット 1-3 | 1                 | No              |
| 6    | 通信命令 2 | 通信パケット 2-3 | 2                 | Yes             |
| 7    | 通信命令 3 | 通信パケット 3-1 | 2                 | No              |
| 8    | 通信命令 1 | 通信パケット 1-4 | 1                 | No              |
| 9    | 通信命令 1 | 通信パケット 1-5 | 1                 | Yes             |
| 10   | 通信命令 3 | 通信パケット 3-2 | 2                 | No              |
| 11   | 通信命令 4 | 通信パケット 4-1 | 1                 | Yes             |
| 12   | 通信命令 3 | 通信パケット 3-3 | 2                 | Yes             |

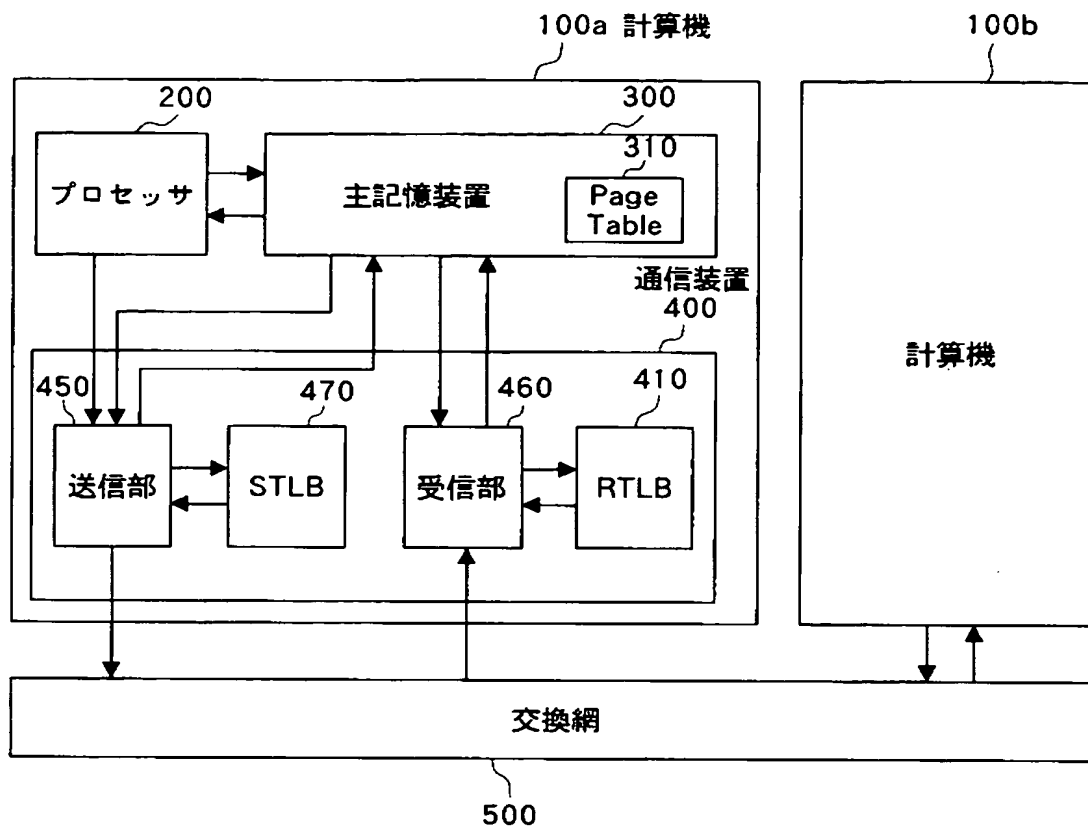
【図 10】



【図 11】

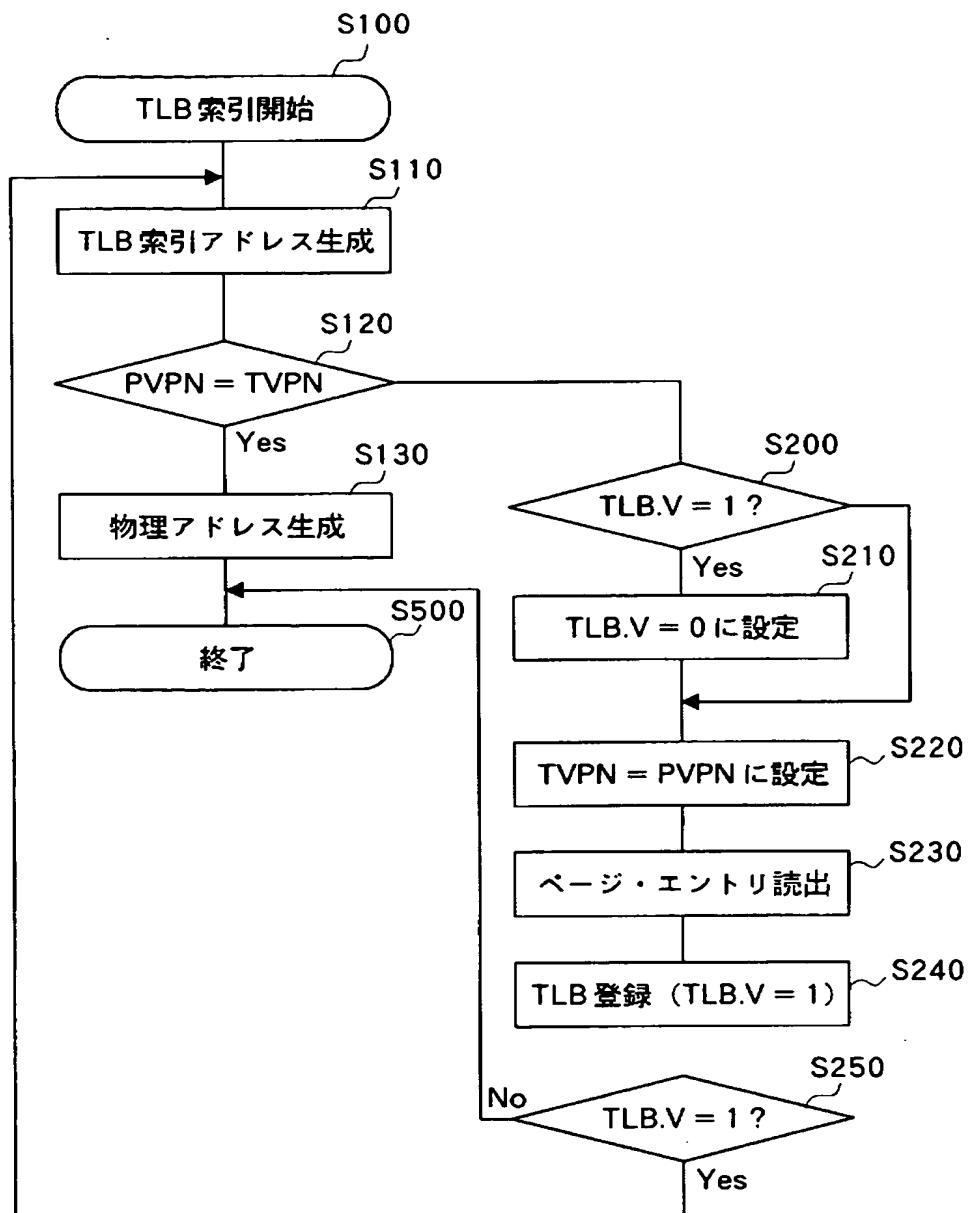


【図 12】





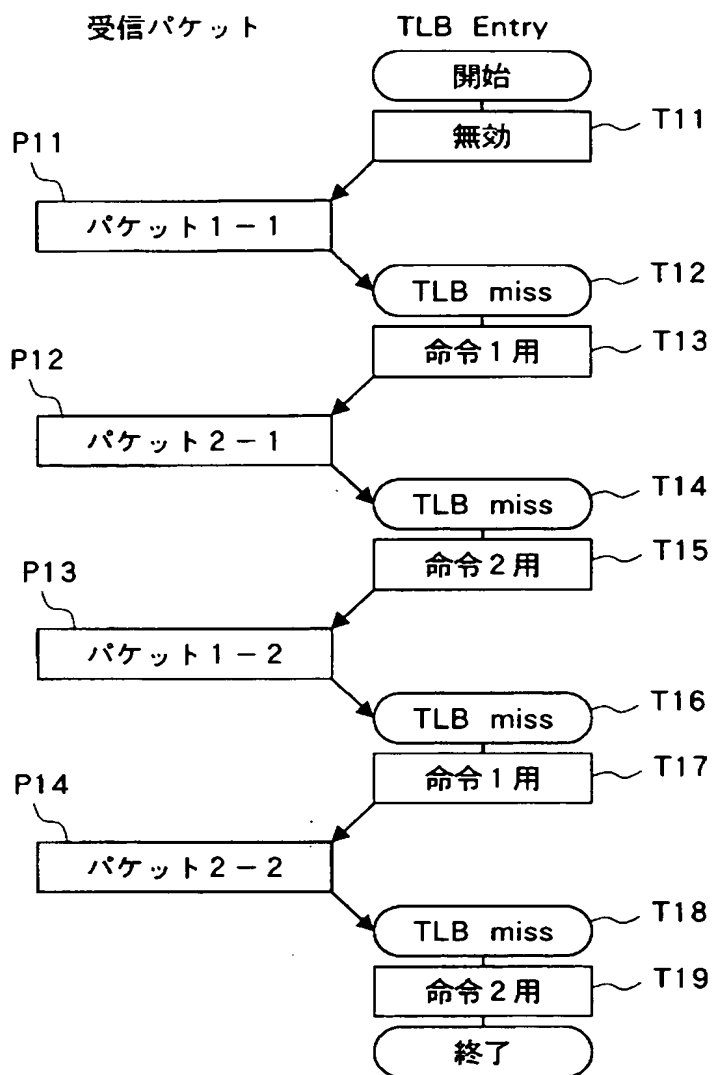
【図 13】



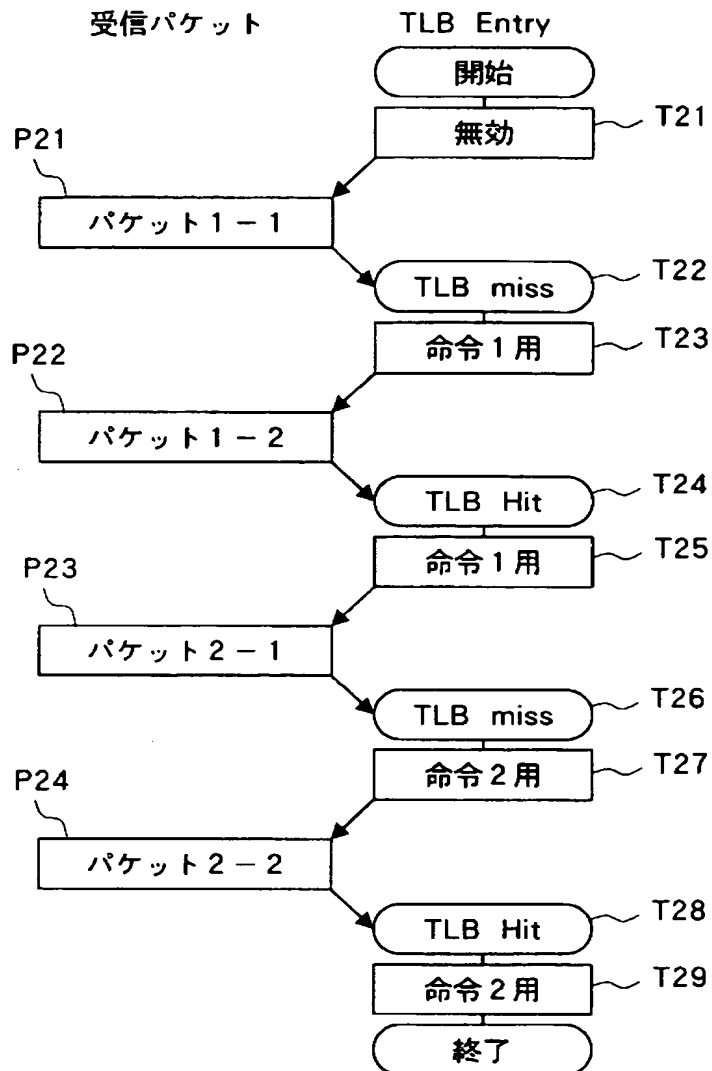
【図 14】

| 送信順序 | 通信命令   | 通信パケット     | 通信パケット内<br>命令識別番号 | 通信命令内<br>最終パケット |
|------|--------|------------|-------------------|-----------------|
| 1    | 通信命令 1 | 通信パケット 1-1 | 1                 | No              |
| 2    |        | 通信パケット 1-2 | 1                 | No              |
| 3    |        | 通信パケット 1-3 | 1                 | No              |
| 4    |        | 通信パケット 1-4 | 1                 | No              |
| 5    |        | 通信パケット 1-5 | 1                 | Yes             |
| 6    | 通信命令 2 | 通信パケット 2-1 | 2                 | No              |
| 7    |        | 通信パケット 2-2 | 2                 | No              |
| 8    |        | 通信パケット 2-3 | 2                 | Yes             |
| 9    | 通信命令 3 | 通信パケット 3-1 | 2                 | No              |
| 10   |        | 通信パケット 3-2 | 2                 | No              |
| 11   |        | 通信パケット 3-3 | 2                 | Yes             |
| 12   | 通信命令 4 | 通信パケット 4-1 | 1                 | Yes             |

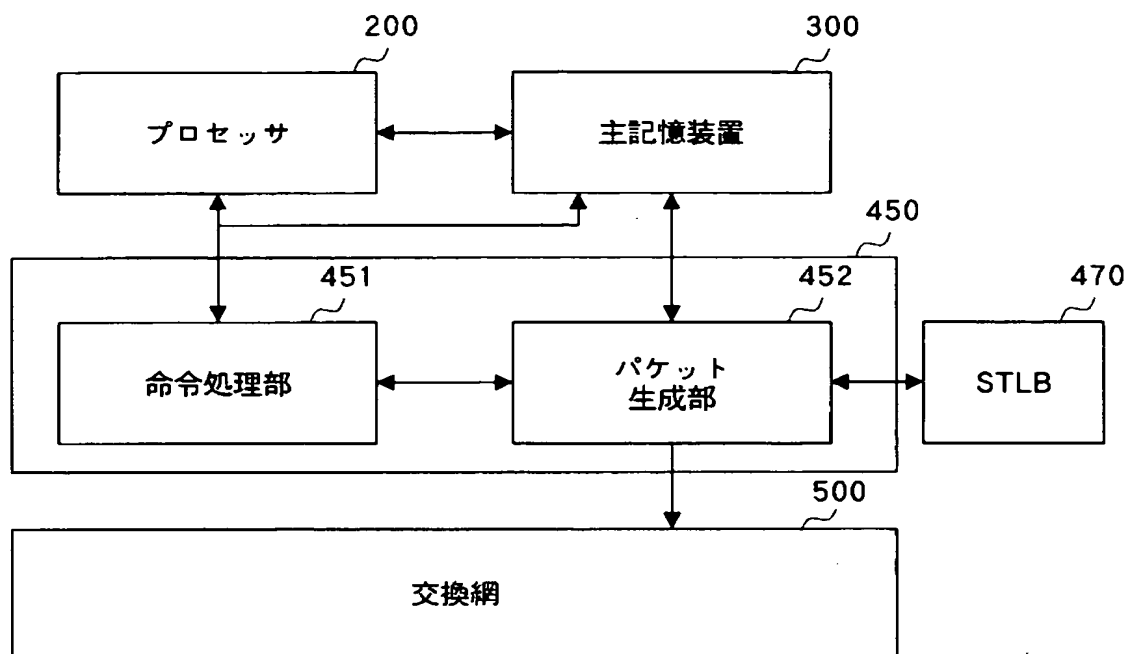
【図 15】



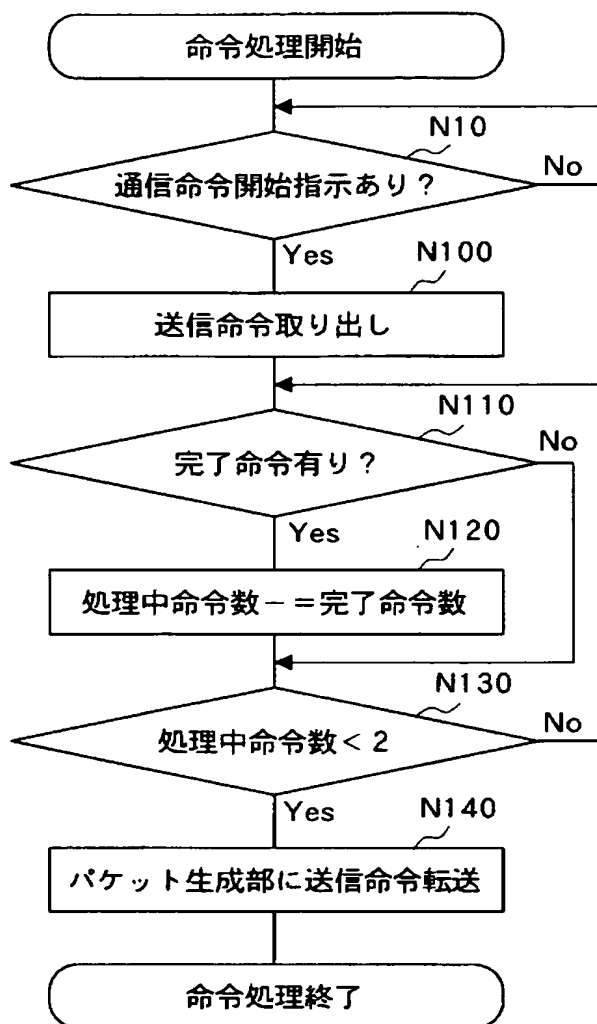
【図 16】



【図 17】



【図 18】



【書類名】 要約書

【要約】

【課題】 T L Bにおいて、変換対象論理アドレスでT L Bエントリを索引する方式ではT L Bヒット率を向上させることが困難である。

【解決手段】 複数の計算機がパケット交換網を介して接続されるネットワークシステムで行なわれるデータ転送方法において、複数の計算機のそれぞれに、主記憶装置と、通信処理命令を発行するプロセッサと、プロセッサからの通信命令を処理し、パケット交換網を介して他の計算機と通信を行う通信装置とを設け、通信装置には、交換網に通信を送信する送信部と交換網から通信を受信する受信部それぞれがT L Bエントリを複数保持するT L Bを設け、プロセッサは送り元の計算機の情報を含む前記通信命令を発行することとし、受信部は、使用するT L Bエントリを送り元の計算機に応じて決定させる。

【選択図】 図6

特願 2 0 0 3 - 0 4 4 1 9 8

出 願 人 履 歴 情 報

識別番号

[ 0 0 0 0 0 4 2 3 7 ]

1. 変更年月日

1 9 9 0 年 8 月 2 9 日

[変更理由]

新規登録

住 所

東京都港区芝五丁目 7 番 1 号

氏 名

日本電気株式会社